

CURVATURES FOR
PARAMETER SUBSETS IN NONLINEAR REGRESSION

R. Dennis Cook
University of Minnesota

and

Miriam L. Goldberg
University of Wisconsin

Technical Report No. 448

March 1985

CURVATURES FOR
PARAMETER SUBSETS IN NONLINEAR REGRESSION

by

R. DENNIS COOK
Department of Applied Statistics
University of Minnesota
St. Paul, Minnesota 55108

and

MIRIAM L. GOLDBERG
Department of Statistics
University of Wisconsin
Madison, Wisconsin 53706

ABSTRACT

The relative curvature measures of nonlinearity proposed by Bates and Watts (1980) are extended to an arbitrary subset of the parameters in a normal, nonlinear regression model. In particular, the subset curvatures proposed indicate the validity of linearization-based approximate confidence intervals for single parameters. The derivation produces the original Bates-Watts measures directly from the likelihood function. When the intrinsic curvature is negligible, the parameter-effects curvature array contains all information necessary to construct curvature measures for parameter subsets.

Key Words: Confidence regions, Curvature measures, Least squares, Likelihood.

1. INTRODUCTION

Confidence regions for parameters of a normal nonlinear regression model are commonly constructed by using linear regression methods, replacing the solution locus with the tangent plane at the maximum likelihood estimate. Such tangent plane regions are generally easier to construct than corresponding likelihood regions. More importantly, the elliptical contours of tangent plane regions are relatively easy to characterize and understand, particularly for one- or two-dimensional parameter subsets which are often of interest. Likelihood regions, on the other hand, are not influenced by parameter-effects nonlinearity and, therefore, generally have true coverage closer to the nominal level than do tangent plane regions. Under suitable regularity conditions and with a sufficiently large sample size, tangent plane and likelihood regions will be in good agreement, but in any particular problem the strength of this agreement is usually uncertain.

Bates and Watts (1980) propose measures of intrinsic and parameter-effects curvature for assessing the adequacy of the tangent plane approximation: Relatively small values for both the maximum intrinsic curvature r^n and the maximum parameter-effects curvature r^T indicate that the tangent plane approximation is reasonable, while relatively large values for either r^n or r^T indicate that this approximation is questionable. These ideas are extended and refined by Bates and Watts (1981), and Hamilton, Bates and Watts (1982). For a review of related literature, see Bates and Watts (1980) and Ratkowsky (1983). Programs for calculating r^n and r^T are given by Bates, Hamilton and Watts (1983).

The material in Bates and Watts (1980) represents an important step forward, but their method for assessing the adequacy of the tangent plane approximation applies only to tangent plane regions for the full parameter vector. This method is not appropriate for assessing the adequacy of tangent plane regions for a subset of parameters, as indicated by Cook and Witmer (1984) and Linssen (1980). It is fairly easy to construct examples where r^T is relatively large and yet there is good agreement between the tangent plane and likelihood regions for a subset of the parameters. One such example is given in Section 2 which is a brief review of the tangent plane approximation and the Bates-Watts methodology. We are often interested in confidence regions for subsets, particularly for individual parameters. Thus, the inability of the Bates-Watts methodology to assess the adequacy of subset regions reflects an important gap in our understanding and ability to deal with nonlinear models.

In Section 3 we develop measures for assessing the agreement between tangent plane and likelihood regions for an arbitrary subset of parameters from a nonlinear regression model. The measures require the same building blocks as needed for the construction of r^T , and reduce to r^T when the full parameter vector is considered. Computationally, these measures require little more effort than r^T itself. Section 4 contains several examples and our concluding comments are given in Section 5. In the remainder of this section, we establish notation and briefly review relevant background information.

A nonlinear regression model can be represented in the form

$$y_i = f(x_i, \theta) + \epsilon_i, \quad i=1, \dots, n \quad (1)$$

where y_i is the i -th response, x_i is a vector of known variables, θ is a $p \times 1$ vector of unknown parameters, the response function f is a known, scalar-valued function that is twice continuously differentiable in θ , and the errors ϵ_i are independent and identically distributed normal random variables with mean 0 and variance σ^2 .

The maximum likelihood (ML) estimator $\hat{\theta}$ of θ can be obtained by minimizing the residual sum of squares

$$RSS(\theta) = \sum_{i=1}^n (y_i - f(x_i, \theta))^2 \quad (2)$$

Kennedy and Gentle (1980) discuss methods for obtaining $\hat{\theta}$. For our purposes we assume that $\hat{\theta}$ is available.

For notational convenience, let $f_i(\theta) = f(x_i, \theta)$ and let V denote the $n \times p$ matrix with elements $f_i^r = \partial f_i / \partial \theta_r$, $i=1, \dots, n$, $r=1, \dots, p$.

Here and in what follows all derivatives are evaluated at $\hat{\theta}$ unless explicitly indicated otherwise.

Various quadratic approximations to be used in the following sections involve the $p \times p$ matrices W_i , $i=1, \dots, n$, with elements $f_i^{rs} = \partial^2 f_i / \partial \theta_r \partial \theta_s$, $r, s=1, \dots, p$. These matrices can be written conveniently in an $n \times p \times p$ array W (Bates and Watts, 1980). The ab -th "column" of W is the ab -th second derivative vector W_{ab} with elements f_i^{ab} , $i=1, \dots, n$, while the i -th face W_i of W is the $p \times p$ matrix consisting of the i -th elements of the second derivative vectors W_{ab} .

2. CURVATURES AND THE TANGENT PLANE APPROXIMATION

Let $F(\theta)$ denote the $n \times 1$ vector with elements $f_i(\theta)$. The standard elliptical confidence region for θ based on replacing $F(\theta)$ with the tangent plane at $\hat{\theta}$ can be written as

$$\{\theta: \phi^T V^T V \phi \leq s^2 G\} \quad (3)$$

where $\phi = (\phi_a) = \theta - \hat{\theta}$, $s^2 = \text{RSS}(\hat{\theta})/(n-p)$, $G = pF_\alpha(p, n-p)$ and $F_\alpha(v_1, v_2)$ is the upper α probability point of an F distribution with v_1 and v_2 degrees of freedom.

To assess the adequacy of the region in (3), we need the standard quadratic expansion of F about $\hat{\theta}$:

$$F(\theta) \approx F(\hat{\theta}) + V\phi + \frac{1}{2} \phi^T W \phi \quad (4)$$

Multiplication involving three-dimensional arrays is defined as in Bates and Watts (1980) so that $\phi^T W \phi$ is an $n \times 1$ vector with elements $\phi^T W_i \phi$, $i=1, \dots, n$. Generally, if F is quadratic over a sufficiently large neighborhood of $\hat{\theta}$ and the quadratic term of (4) is sufficiently small relative to the linear term, the tangent plane region (3) should be reasonable; otherwise, this approximation may be in doubt. Bates and Watts (1980, 1981) implement this idea by first decomposing each column of W into its projections onto the column and null spaces of V : $W_{ab} = P_V W_{ab} + (I - P_V) W_{ab} = W_{ab}^T + W_{ab}^N$, where P_V is the orthogonal projection operator for the column space of V . With this decomposition, the quadratic expansion (4) becomes

$$F(\theta) \approx F(\hat{\theta}) + V\phi + \frac{1}{2} \phi^T W^\tau \phi + \frac{1}{2} \phi^T W^\eta \phi \quad (5)$$

where W^τ and W^η are the $n \times p \times p$ arrays whose columns are W_{ab}^τ and W_{ab}^η , respectively.

Next, the adequacy of the tangent plane region is assessed by using the maximum parameter-effects curvature

$$\Gamma^\tau = \max \frac{||\phi^T W^\tau \phi||}{||V\phi||^2} \sqrt{p_s} \quad (6)$$

and the maximum intrinsic curvature

$$\Gamma^\eta = \max \frac{||\phi^T W^\eta \phi||}{||V\phi||^2} \sqrt{p_s} \quad (7)$$

where the maximum is taken over all ϕ in R^p . These curvatures as well as the decomposition of $\phi^T W \phi$ displayed in (5), reflect different characteristics of the nonlinearity of the model. The intrinsic curvature Γ^η is invariant under reparameterizations and is thus a measure of the intrinsic nonlinearity of the solution locus. In contrast, Γ^τ depends on the parameterization: different parameterizations can result in substantially different values of Γ^τ . If both Γ^τ and Γ^η are sufficiently small, the tangent plane region (3) should be adequate.

More specifically, for a tangent plane region of the form (3), Bates and Watts (1980) suggest that the linear approximation should be adequate if Γ^η and Γ^τ are both small compared to the guide $c = 1/\sqrt{F_\alpha(p, n-p)}$. When Γ^η or Γ^τ is greater than c , the linear approximation and the circular approximation that is the basis of the curvature measures both break down within the tangent

plane region. Thus, Ratkowsky (1983) proposes that $c/2$ be used as a cutoff level, beyond which the tangent plane region is presumed inadequate.

To demonstrate that the Bates-Watts methodology can fail for subsets of θ , we consider the Fieller-Creasy problem in which the ratio of the means of two normal populations is of interest. The corresponding nonlinear model can be written as

$$f(x_1, \theta) = \theta_1 x_1 + \theta_1 \theta_2 (1 - x_1) \quad (8)$$

where x_1 is an indicator variable that takes the values 1 and 0 for populations 1 and 2, respectively. For convenience we assume equal sample sizes for the two populations $n_1 = n_2 = n/2$ and, without loss of generality, we assume that σ^2 is known.

The model given in (8) is intrinsically linear so that $r^n = 0$. Further, Cook and Witmer (1984) show that

$$r^T = \frac{\sqrt{2}\sigma\{(\hat{\theta}_2^2 + 1)^{1/2} + |\hat{\theta}_2|\}}{|\hat{\theta}_1| \sqrt{n}} \quad (9)$$

In this case the Bates-Watts (1980) guide for judging the adequacy of the tangent plane approximation is $c = (\chi(\alpha; 2))^{1/2}$ where $\chi(\alpha; v)$ is the upper α probability point of the chi-squared distribution with v degrees of freedom. However, it is clear that standard methods can be used to form exact confidence intervals for θ_1 , the mean of the first population, regardless of the value of r^T . In other words, the tangent plane and likelihood regions for θ_1 are identical for all r^T .

A similar phenomenon occurs in connection with θ_2 . Let $r = \sigma^2 \chi(\alpha; 1) / n \hat{\theta}_1^2$.

Assuming that $r < 1$, Cook and Witmer (1984) show that the $1-\alpha$ likelihood region for θ_2 can be written as

$$[\hat{\theta}_2 \pm \{r(1-r) + r\hat{\theta}_2^2\}^{1/2}]/(1-r) \quad (10)$$

The level associated with this region is exact. The corresponding tangent plane region is

$$\hat{\theta}_2 \pm (r + r\hat{\theta}_2^2)^{1/2} \quad (11)$$

Clearly, (10) and (11) will be close only if r is sufficiently small. For any fixed value of r , however, r^T may be large or small depending on the value of $\hat{\theta}_2$ so that again the Bates-Watts criterion fails to reflect accurately the agreement between the tangent plane and likelihood regions for a parameter subset. We will return to this example at the end of the next section.

3. SUBSETS

Let $L(\theta, \sigma^2)$ denote the log likelihood for model (1), and partition $\theta^T = (\theta_1^T, \theta_2^T)$ where θ_1 is a $p_1 \times 1$ vector, $i=1,2$. The standard likelihood region for θ_2 can be written in the form (Cox and Hinkley, 1974, p. 343).

$$\{\theta_2: 2[L(\hat{\theta}, \hat{\sigma}^2) - L(g(\theta_2), \theta_2, \tilde{\sigma}^2(\theta_2))] \leq \rho\} \quad (12)$$

where ρ , a selected positive constant, is used to set the nominal level and $(g^T(\theta_2), \tilde{\sigma}^2(\theta_2))$ represents the vector-valued function that maximizes

$L(\theta_1, \theta_2, \sigma^2)$ for each value of θ_2 . Evaluating (12), the likelihood region for θ_2 can be written equivalently as

$$\{\theta_2 : n \cdot \log \left[\sum_{i=1}^n (y_i - f_i(g(\theta_2), \theta_2))^2 / n \hat{\sigma}^2 \right] \leq p\} \quad (13)$$

Clearly, the form of this region is governed by the vector-valued function $h(\theta_2) = F(g(\theta_2), \theta_2)$. If h is essentially linear over a sufficiently large neighborhood of $\hat{\theta}_2$, the contours of (13) will be elliptical and we can expect (13) and the corresponding tangent plane region to agree; otherwise these regions will tend to be dissimilar. To determine when these regions are in substantial agreement, we investigate the behavior of h by using the method described in Section 2, except that F is replaced by h which, in combination with $Y = (y_i)$, contains essential information on θ_2 . Thus, in exact analogy with the Bates-Watts development, we will produce expressions for the curvature of the solution locus submanifold defined by h . Where necessary for clarity, we refer to this as "subset curvature". Similarly, "subset parameter-effects", and "subset intrinsic" refer to the decomposition of the subset curvature into components in the submanifold tangent plane and its orthogonal complement.

Let $\alpha^T(\theta_2) = (\alpha_1(\theta_2)) = (g^T(\theta_2), \theta_2^T)$, let Δ_1 denote the $p \times p_2$ matrix with elements $\partial \alpha_i / \partial \theta_{2j}$, $i=1,2,\dots,p$, $j=1,2,\dots,p_2$, and let Δ_2 denote the $p \times p_2 \times p_2$ array with i -th face Δ_{2i} , $i=1,2,\dots,p$; the elements of Δ_{2i} are $\partial^2 \alpha_i / \partial \theta_{2j} \partial \theta_{2k}$, $j,k=1,\dots,p_2$. We assume, of course, that g is a twice continuously differentiable function of θ_2 . With these definitions the straightforward quadratic approximation of $h(\theta_2)$ about $\hat{\theta}_2$ can be written as

$$h(\theta_2) \approx F(\hat{\theta}) + V\Delta_1\phi_2 \quad (14a)$$

$$+ \frac{1}{2}\phi_2^T \Delta_1^T W \Delta_1 \phi_2 \quad (14b) \quad (14)$$

$$+ \frac{1}{2}V(\phi_2^T \Delta_2 \phi_2) \quad (14c)$$

where $\phi_2 = \theta_2 - \hat{\theta}_2$.

3.1 Refining Equation (14).

For the quadratic expansion in (14) to be useful, we need to develop explicit forms for Δ_1 and Δ_2 to produce a reexpression of (14) that displays the (subset) parameter-effects and intrinsic components of h at $\hat{\theta}_2$. To avoid interruption, the details of this development have been relegated to the Appendix. Here we discuss the final form.

The final form of (14) is based on the assumption that the intrinsic curvature of F at $\hat{\theta}$ is negligible. That assumption is somewhat restrictive but it is valid in the important class of problems where the parameters of interest are nonlinear functions of the location parameters in a linear model. In any event, we judge the practical advantages of allowing for substantial intrinsic curvatures to be minimal since experience has shown (See Bates and Watts 1980, and Ratkowsky 1983) that they are typically small. Of course, r^η can and should be evaluated in practice so that this assumption can be checked.

In the remainder of this paper we use $C(M)$ and $C'(M)$ to indicate the column and null spaces, respectively, of the matrix M ; the corresponding orthogonal projection operators will be denoted by P_M and P_M' , respectively.

In their development of the intrinsic and parameter-effects curvatures for

the full parameter vector, Bates and Watts (1980) found it convenient and revealing to work in transformed coordinates. Similarly, the quadratic expansion (14) is most easily understood in terms of these same transformed coordinates: Let $V = UR$ denote the unique QR-factorization of V where R is upper triangular and the columns of the $n \times p$ matrix U form an orthonormal basis for $C(V)$. Next, partition R as

$$R = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \quad (15)$$

where R_{1i} is $p_i \times p_i$, $i=1,2$. Transformed coordinates $\tilde{\phi}$ can now be defined as $\tilde{\phi}^T = (\tilde{\phi}_1^T, \tilde{\phi}_2^T) = \phi^T R^T$ so that

$$\tilde{\phi}_1 = R_{11}\phi_1 + R_{12}\phi_2 \quad (16)$$

and

$$\tilde{\phi}_2 = R_{22}\phi_2 \quad (17)$$

In the following any quantity with a tilde added above indicates evaluation in the $\tilde{\phi}$ coordinates. Thus, for example,

$\tilde{V} = U$ and $\tilde{W} = R^{*T} W R^{*1}$. Partition the i th face \tilde{W}_i of \tilde{W} as

$$\tilde{W}_i = \begin{pmatrix} \tilde{W}_{i11} & \tilde{W}_{i12} \\ \tilde{W}_{i21} & \tilde{W}_{i22} \end{pmatrix}, \quad i=1, \dots, n \quad (18)$$

where the dimension of \tilde{W}_{ijj} is $p_j \times p_j$, $j=1,2$. Next, define \tilde{W}_{22} to be the

$n \times p_2 \times p_2$ subarray of \tilde{W} with i -th face \tilde{W}_{i22} and similarly define \tilde{W}_{i12} to be the $n \times p_1 \times p_2$ subarray of \tilde{W} with i -th face \tilde{W}_{i12} , $i=1, \dots, n$. Finally, partition $V = (V_1, V_2)$ and $U = (U_1, U_2)$ where U_1 and V_1 are $n \times p_1$ matrices.

With this structure, the quadratic expansion of h can be reexpressed informatively as

$$h(\theta_2) \approx F(\hat{\theta}) + U_2 \tilde{\phi}_2 \quad (19a)$$

$$+ \frac{1}{2} \tilde{\phi}_2^T [P_{U_2}] [\tilde{W}_{22}] \tilde{\phi}_2 \quad (19b) \quad (19)$$

$$= U_1 [\tilde{\phi}_2^T U_2^T] [\tilde{W}_{12}] \tilde{\phi}_2 \quad (19c)$$

where the brackets $[\cdot][\cdot]$ indicate column (sample space) multiplication as defined in Bates and Watts (1980), and discussed briefly in the Appendix. Term (19a) describes the plane tangent to h at $\hat{\theta}_2$. Since $C(U_2) = C(P_V' V_2)$, this plane is simply the affine subspace $F(\hat{\theta}) + C(P_V' V_2)$. This is the same as the subspace obtained when using the tangent plane approximation to form a confidence region for θ_2 . In other words, the confidence contour based on the tangent plane approximation will coincide with those based on substituting the linear approximation of h into (13), as expected.

Term (19b) contains the projections of the columns of \tilde{W}_{22} onto the plane tangent to h at $\hat{\theta}_2$. Thus, this term reflects the (subset) parameter-effects curvature of h in the direction $\tilde{\phi}_2$. The maximum parameter-effects curvature Γ_s^T for the subset θ_2 can now be defined as

$$r_s^T(\theta_2) = \max_{d \in D} ||d^T [P_{U_2}] [\tilde{W}_{22}] d|| \sqrt{p_2} s \quad (20)$$

where the maximum is taken over all d in $D = \{d: d \in \mathbb{R}^2, ||d|| = 1\}$. Since $\tilde{\phi}_2$ is a linear transformation of ϕ_2 as described in (17), $r_s^T(\theta_2)$ will be the same in both coordinate systems:

To further understand (20), partition the i th face A_i of the $p \times p \times p$ unscaled parameter-effects curvature array $A = [U^T][\tilde{W}]$ as

$$A_i = \begin{pmatrix} A_{i11} & A_{i12} \\ A_{i21} & A_{i22} \end{pmatrix} \quad (21)$$

where the dimension of A_{ijj} is $p_j \times p_j$, $j=1,2$, $i=1, \dots, p$. Next, let A_{22} denote the $p_2 \times p_2 \times p_2$ subarray of A with faces A_{i22} , $i=p_1+1, \dots, p$. Then

$$[P_{U_2}][\tilde{W}_{22}] = [U_2][A_{22}]$$

and

$$r_s^T(\theta_2) = \max_D ||d^T A_{22} d|| \sqrt{p_2} s \quad (22)$$

In this form it is clear that the maximum parameter-effects curvature for the subset problem depends only on the behavior of the $\tilde{\phi}_2$ parameter-curves. The elements of A_{22} can be used to understand the behavior of these parameter-curves in terms of arcing, "compansion", fanning and torsion, as described in Bates and Watts (1981).

Term (19c) is clearly in $C(V_1)$ and is thus orthogonal to the subspace tangent plane. This term then reflects the intrinsic curvature of h at $\hat{\theta}_2$ so

that the maximum intrinsic curvature can be defined as

$$r_s^\eta(\theta_2) = \max_D ||[d^T U_2^T][\tilde{W}_{12}]d|| 2\sqrt{p_2}s \quad (23)$$

Note that (23) contains the extra factor 2, corresponding to the absence of the factor 1/2 in (19c).

This curvature can also be expressed in terms of a subarray of A. Let A_{12} denote the $p_2 \times p_1 \times p_2$ subarray of A that has faces A_{i12} , $i=p_1+1, \dots, p$. Then $A_{12} = [U_2^T][\tilde{W}_{12}]$ and

$$\begin{aligned} r_s^\eta(\theta_2) &= \max_D ||[d^T][A_{12}]d|| 2\sqrt{p_2}s \\ &= \max_D || \sum_{j=p_1+1}^p d_j A_{j12} d || 2\sqrt{p_2}s \end{aligned} \quad (24)$$

where d_j is the $(j-p_1)$ th element of d . Interestingly, the intrinsic curvature for the subset problem depends only on fanning and torsion components of A; compansion and arcing play no role in the determination of r_s^η . The fanning and torsion terms of A depend in part on how the columns of V are ordered. Since we have assumed that the last p_2 columns of V correspond to θ_2 , it is the fanning and torsion with respect to this ordering that are important.

If both r_s^η and r_s^τ are sufficiently small, the likelihood and tangent plane confidence regions for θ_2 will be similar; otherwise we can expect these regions to be dissimilar. Following Bates and Watts (1980), $c = (F_\alpha(p_2, n-p))^{-1/2}$ can be used as a rough guide for judging the size of these curvatures. As noted earlier, our experience indicates that curvatures must be substantially

less than c to insure close agreement between tangent plane and likelihood regions. This will be illustrated in sections 3.3 and 4.

Finally, we combine the intrinsic and parameter-effects components of (19) to define the total curvature $\Gamma_s(\theta_2)$ of h at $\hat{\theta}_2$ as

$$\Gamma_s(\theta_2) = \sqrt{p_2} s \max_D \{ ||d^T A_{22} d||^2 + 4 ||[d^T] [A_{12}] d||^2 \}^{1/2} \quad (25)$$

As will be demonstrated in the next subsection, the total subset curvature Γ_s may be more relevant than both Γ_s^η and Γ_s^τ . For example, it is possible to have $\Gamma_s^\eta < c$ and $\Gamma_s^\tau < c$ while $\Gamma_s > c$. In such situations Γ_s^τ and Γ_s^η may incorrectly indicate that the tangent plane approximation is adequate, while Γ_s correctly indicates otherwise.

When the full parameter θ is of interest, we have $\theta_2 = \theta$ and $p_2 = p$. In this case, the subset intrinsic curvature (24) is zero, A_{22} is the Bates-Watts parameter-effects array, and both (22) and (25) represent the maximum parameter-effects curvature for θ . Thus, our derivation based on the likelihood reproduces the primary quantity developed by Bates and Watts (1980).

The main conclusion of this section is that the unscaled parameter-effects curvature array A for the full parameter contains all necessary information for evaluating the adequacy of tangent plane confidence regions for certain subsets of θ . For example, if the last parameter θ_p is of interest then $\Gamma_s^\tau(\theta_p)$ is simply $s |a_{pp}|$ where a_{ijk} is the (j,k) -th element of the i -th face of A . Similarly,

$$\Gamma_s^n(\theta_p) = 2s \left(\sum_{i=1}^{p+1} a_{p+1}^2 \right)^{1/2} \quad (26)$$

Thus, compansion and fanning are the only effects that are relevant to an assessment of the agreement between likelihood and tangent plane confidence regions for a single parameter.

3.2 Computation

Recall that the developments of this section are based on the assumption that the last p_2 columns of V correspond to the parameters of interest. This assumption is necessary to maintain the collective identity of θ_2 as indicated in (17). This implies that the ordering of the columns of V is critical and consequently θ_p is the only single parameter for which curvatures can be constructed from a given parameter-effects array A . The A -array for other orderings can be constructed by permuting the columns of V and beginning again, of course.

Alternatively, a computationally more efficient method for obtaining the A -array in a rotated coordinate system can be constructed as follows. Let $\phi_z = Z\phi$ where Z is a selected $p \times p$ permutation matrix. In what follows, the subscript z added to any quantity indicate evaluation in the coordinates ϕ_z . Clearly, $V_z = VZ^T = URZ^T$. Let U^{*T} be an orthogonal matrix such that $R^* = U^{*T}RZ^T$ is upper triangular. Since the QR-factorization of V_z is unique, it follows that $V_z = U_z R_z$ where $U_z = U U^{*T}$ and $R_z = R^*$. Using this structure it is not difficult to verify that

$$A_z = [U^*][U^*AU^{*T}] \quad (27)$$

Thus, to find A_z , the parameter-effects curvature array for the rotated coordinates ϕ_z , we need only the $p \times p$ matrix U^* to diagonalize RZ^T . A single call to LINPACK (1979) routine SCHEX produces R^* , $[U^{*T}][A]$ and the information necessary to construct U^* .

3.3 Fieller-Creasy Again

To apply Γ_S^n and Γ_S^T in the Fieller-Creasy problem when θ_2 is the subset of interest, we require only the 2×2 parameter-effects curvature array A for

$$V = (x + \hat{\theta}_2(b-x), \hat{\theta}_1(b-x))$$

where x is the $n \times 1$ vector with elements x_i as defined following (8) and b is an $n \times 1$ vector of one's. The faces A_i of A are (Cook and Witmer, 1984)

$$A_1 = \frac{\hat{\theta}_2 \sqrt{2}}{\hat{\theta}_1 \{n(1+\hat{\theta}_2)\}^{1/2}} \begin{bmatrix} 0 & 1 \\ 1 & -2\hat{\theta}_2 \end{bmatrix} \quad (28)$$

and

$$A_2 = A_1 / \hat{\theta}_2 \quad (29)$$

Reading directly from this array we have

$$\Gamma_s^T(\theta_2) = \sigma |a_{222}| = \frac{2^{3/2}\sigma}{\sqrt{n}|\hat{\theta}_1|} \cdot \frac{|\hat{\theta}_2|}{(1 + \hat{\theta}_2^2)^{1/2}} \quad (30)$$

and

$$\Gamma_s^\eta(\theta_2) = 2\sigma |a_{212}| = \frac{2^{3/2}\sigma}{\sqrt{n}|\hat{\theta}_1|} \cdot \frac{1}{(1 + \hat{\theta}_2^2)^{1/2}} \quad (31)$$

Recall that we are assuming σ to be known in this example so that the guide for assessing the magnitudes of Γ_s^η and Γ_s^T is $c = (\chi(\alpha; 1))^{1/2}$.

From (30) we see that $\Gamma_s^T(\theta_2)$ will be zero only if $\hat{\theta}_2 = 0$; in this case $\Gamma_s^\eta(\theta_2) = 2^{3/2}\sigma/\sqrt{n}|\hat{\theta}_1| < c$ or, equivalently, $r = 2\sigma^2\chi(\alpha; 1)/n\hat{\theta}_1^2 < 1/4$ is necessary for the subset intrinsic curvature to be less than the guide. Further $r < 1/4$ is a sufficient although not necessary condition for both $\Gamma_s^\eta(\theta_2)$ and $\Gamma_s^T(\theta_2)$ to be less than c when $\hat{\theta}_2$ is arbitrary.

Next, using (25) it follows that the total subset curvature is simply

$$\Gamma_s(\theta_2) = 2^{3/2}\sigma/\sqrt{n}|\hat{\theta}_1| \quad (32)$$

and thus $\Gamma_s(\theta_2) < c$ if and only if $r < 1/4$. When $r > 1$ the likelihood region for θ_2 will be either the complement of an interval or else the entire real line; otherwise this region will be the interval given in (10). In this example, the total subset curvature recovers the critical quantity r as introduced in section 2, and the condition $\Gamma_s < c$ insures that the tangent plane interval (11) will in fact be approximating a likelihood interval rather than some dissimilar region. This condition also provides for an added measure of agreement between these intervals since it is equivalent to $r < 1/4$ rather than simply $r < 1$.

Applying (22) and (24) when θ_1 is the subset of interest gives $r_s^n(\theta_1) = r_s^T(\theta_1) = 0$, as expected. Notice that this conclusion cannot be obtained by inspecting the A array given in (28) and (29). As mentioned previously, different subsets in general require different orderings for the columns of V and thus different coordinates. This is the case here.

Finally, we consider the special case characterized by $(\hat{\theta}_1, \hat{\theta}_2) = (3, 0)$ and $r = .428$. These conditions correspond to $n = 2\sigma^2$. From (9), $r^T = .33 < .41 = \chi^{.41/2}(.05; 2)$. From Figure 1 (Cook and Witmer 1984), we see that the likelihood region, whose level is exact in this case, does not seem to be adequately approximated by the tangent plane region for small values of θ_1 .

Further insight into this problem can be gained by inspecting marginal regions for θ_1 and θ_2 . Generally, marginal regions for subsets can be obtained by projecting all points in the joint region onto the appropriate subspaces. The projections of the regions in Figure 1 onto the θ_1 axis show that the likelihood and tangent plane intervals for θ_1 will be identical, as expected. The projections onto the θ_2 axis show that the resulting 98.6 percent likelihood interval will be about 60 percent longer than the corresponding tangent plane interval! This dissimilarity is clearly indicated by $r_s^n(\theta_2) = .67 > .41 = \chi^{.41/2}(.014; 1)$.

Our experience leads to the following heuristic characterization of the problem described in the previous paragraph. Consider a p_2 -dimensional subset θ_2 with guide $c_2 = (F_\alpha(p_2, n-p))^{.41/2}$ and partition $\theta_2^T = (\theta_{21}^T, \theta_{22}^T)$ where θ_{21} is $p_{21} \times 1$, $i=1, 2$. The guide corresponding to the confidence region for θ_{21} obtained by projecting the selected $1-\alpha$ region for θ_2 is simply

$c_{2i} = c_2(p_{2i}/p_2)^{1/2}$, $i=1,2$. When the subset curvatures for θ_{21} are large relative to c_{21} and the subset curvatures for θ_{22} are near zero, it can happen that the curvatures for θ_2 are moderate. In such cases the curvatures for θ_2 can provide a misleading indication that the tangent plane and likelihood regions for θ_2 are in acceptable agreement. As hinted above, this problem might be overcome by requiring that all subsets θ_{21} of θ_2 have curvatures less than the respective guides c_{21} . When $\theta_2 = \theta$ this added requirement seems to represent a useful fine tuning of the basic Bates-Watts methodology.

4. ILLUSTRATIONS

In this section we present several numerical examples to illustrate selected results of the previous sections.

For the first example we use the Michaelis-Menton model

$$f_i = \theta_1 x_i / (\theta_2 + x_i) \quad (33)$$

in combination with the 12 observations reported in Bates and Watts (1980). Figure 2 gives 87 percent tangent plane (broken contour) and likelihood (solid contour) confidence regions for (θ_1, θ_2) . Here and in the following examples the levels of displayed bivariate confidence regions are chosen so that the corresponding univariate marginal regions have a nominal 95 percent coverage rate. It seems clear from Figure 2 that the tangent plane region for (θ_1, θ_2) is not an adequate approximation of the likelihood region, although interpreting the Bates-Watts guide directly as the cutoff value would lead to the opposite conclusion, since $r^T = .598 < c = .635$. The subset curvatures for θ_1 and θ_2 are listed in Table 1; the corresponding guide is $c = .449$. Again, the curvatures are less than the guide while the marginal likelihood regions do not seem to be well represented by the corresponding tangent plane

regions. This reinforces our previous remark that curvatures must be substantially less than c to insure close agreement. With this interpretation we see that all curvatures successfully indicate the dissimilarity between the various likelihood and tangent plane regions in Figure 2.

Figure 3 gives 88% likelihood and tangent plane regions for (θ_1, θ_2) obtained by using model (33) and the 7 observations reported by Michaelis and Menton (1913). For these data $\Gamma^T = .079$. This value and the subset curvatures reported in Table 1 are relatively small, indicating reasonable agreement between the regions displayed in Figure 3.

For our next example we use the exponential model

$$f_1 = \theta_1 (1 - \exp(\theta_2 x_1)) \quad (34)$$

in combination with the 6 observations reported in Draper and Smith (1981, p. 522., data set 3). In this case $\Gamma^T = 1.92$ clearly indicates the dissimilarity between the 88 percent regions for (θ_1, θ_2) shown in Figure 4. However, the 95% marginal regions for θ_2 are in close agreement, while the agreement between the marginal regions for θ_1 seems less than adequate. These conclusions are clearly indicated by the subset curvatures $\Gamma_s(\theta_2) = .069$ and $\Gamma_s(\theta_1) = .314$ which may be judged relative to the guide $c = .360$.

For the three-parameter asymptotic regression model

$$f_1 = \theta_1 + \theta_2 \exp(\theta_3 x_1) \quad (35)$$

and the 27 observations reported in Ratkowsky (1983, p. 101, data set 1), we obtain $\Gamma^T = 1.53$. The corresponding guide is $c = [F_{.05}(3, 24)]^{1/2} = .58$. This suggests that the 95 percent likelihood region for $\theta^T = (\theta_1, \theta_2, \theta_3)$ cannot be adequately approximated by the corresponding tangent plane region. The subset curvatures for selected subsets of θ are listed in Table 1. From these curvatures alone we would reach the following conclusions: 1) The likelihood

and tangent plane regions for θ_2 are in very close agreement. 2) The marginal regions for θ_1 and θ_3 will be noticeably different, but the agreement is probably adequate for most purposes. 3) The usual 95 percent tangent plane regions for (θ_1, θ_3) and (θ_2, θ_3) should be used for only very rough analyses, although lower level regions may be acceptable replacements for the corresponding likelihood regions. These conclusions are supported by the 86 percent regions for (θ_2, θ_3) and (θ_1, θ_3) shown in Figures 5 and 6, respectively.

For our final example we again use the asymptotic regression model (35), this time in combination with the 9 observations reported by Hunt (1970). Subset curvatures for 4 parameter subsets are listed in Table 1. The subset curvature for θ_3 is small, indicating good agreement between the corresponding likelihood and tangent plane regions. The subset curvatures for the remaining subsets, particularly (θ_2, θ_3) , are large.

The 87 percent likelihood and tangent plane confidence regions for (θ_2, θ_3) are given in Figure 7. The large total curvature, $\Gamma_s(\theta_2, \theta_3) = 36.4$, correctly indicates that use of the tangent plane region as an approximation of the disjoint likelihood region would be a disaster for this pair of parameters. In fairness, however, it should be recalled that the approximations used to derive the subset curvatures are local so that $\Gamma_s(\theta_1, \theta_2)$ is responding primarily to the disagreement between the tangent plane region and the portion of the likelihood region that contains $\hat{\theta}$. Similar comments apply when only θ_2 is of interest.

From Figure 7, there is reasonable agreement between the tangent plane and likelihood regions for θ_3 , as indicated by the small curvature $\Gamma_s(\theta_3) = .095$.

It can be argued justifiably, however, that this correct indication from the curvature is largely fortuitous since the curvatures do not recognize the contribution of the smaller piece of the likelihood region for (θ_2, θ_3) to the likelihood region for θ_3 . Under this argument, the subset curvature measure for θ_3 has failed to indicate the dissimilarity between the tangent plane region for θ_3 and the likelihood region $(-.0191, 0)$ obtained by using only the larger subregion that contains $\hat{\theta}$.

The reason that the curvatures give some inappropriate indications in this final example is that both the linear and quadratic approximations to the model function fail. This failure is evident from a very low R^2 from the regression used by Goldberg, Bates and Watts (1983) to obtain numerical curvatures, and from related measures of "lack of quadraticity" explored by the present authors. In cases where the quadratic approximation to the model function is poor, curvature measures based on that approximation may not be meaningful.

Nevertheless, these subset curvature measures represent an important advance in our understanding of nonlinear models, and provide useful information about the adequacy of the linear approximation when the quadratic approximation is appropriate. Further work is needed on methods of identifying cases where the quadratic approximation may fail.

5. CONCLUSIONS

The subset curvatures developed in this paper appear to be reliable indicators of the adequacy of tangent plane confidence regions for most nonlinear models. In particular, the curvature for a single parameter is a useful tool for assessing the agreement between standard large sample

confidence intervals and corresponding marginal likelihood regions. This ability to deal with subsets greatly extends the usefulness of the Bates-Watts methodology.

Because the original Bates-Watts framework applies only to the complete parameter vector, guidelines developed in that framework can be misleading when the adequacy of the linear approximation is very different for different subsets. To ensure good agreement between the tangent plane and likelihood regions, the maximum curvature must be considerably smaller than the Bates-Watts guide. However, this criterion can be too stringent for certain parameter subsets if the whole-parameter curvature Γ^T is used. By contrast, the subset curvature describes the shape of the likelihood region in the parameter subspace of interest. Thus, the subset curvature is more directly relevant to the tangent plane adequacy question and, based on the examples described above, is evidently more accurate.

The practical usefulness of the methods described here depends, in part, on their ease of implementation. The subset curvatures for any selected subset can be computed directly from the Bates-Watts parameter-effects curvature array. This array can be obtained either analytically (Bates and Watts, 1980) or numerically by using the procedure given in Goldberg, Bates and Watts (1983).

The usefulness of the subset curvatures depends also on the restriction that the intrinsic curvature of F at $\hat{\theta}$ is small. This restriction is not of great practical importance since it has been found to hold in most cases. Nevertheless, a unified approach which incorporates the intrinsic curvature component might offer further insight in some situations.

Another area for further research is the development of measures that

indicate when the subset curvatures themselves may be unreliable due to the failure the second-order approximation to the model function. While the possibility of such failure is of concern, the class of models adequately described by a quadratic function is considerably larger than the class for which the linear approximation alone is adequate.

ACKNOWLEDGEMENTS

Data computations for this work were performed on the University of Wisconsin Statistics Department's research computer. The authors thank Douglas Bates for access to software and data libraries used for some examples presented, and for suggestions on computational methods.

Derivation of Equation (19)

To develop equation (19) from equation (14), we first require explicit expressions for Δ_1 and Δ_2 .

A.1. Δ_1 and Δ_2

Let \ddot{L} and \ddot{L} denote the $p \times p$ matrix and $p \times p \times p$ array of second and third partial derivatives of the log likelihood L with respect to the elements of θ , respectively. Let g_a denote the a -th component of g as defined following (12) and partition \ddot{L} as

$$\ddot{L} = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}$$

where L_{jj} is $p_j \times p_j$, $j=1,2$.

Since g maximizes $L(\theta_1, \theta_2)$ for each fixed value of θ_2 we clearly have

$$\left. \frac{\partial L(g(\theta_2), \theta_2)}{\partial g_a} \right|_{g=g(\theta_2)} = 0 \quad (A.1)$$

for $a=1,2,\dots,p_1$ and all θ_2 . This identity will be used as the basis for obtaining Δ_1 and Δ_2 .

Differentiating both sides of (A.1) with respect to θ_2 and evaluating at

$\hat{\theta}_2$ gives

$$(L_{11}, L_{12})\Delta_1 = 0$$

Since the submatrix consisting of the last p_2 rows of Δ_1 is an identity matrix it follows that

$$\Delta_1 = \begin{pmatrix} {}^{\#}L_{11}^{\#1} & L_{12} \\ I \end{pmatrix} \quad (A.2)$$

Let $e_i = y_i = f_i(\hat{\theta})$. The the first term of

$$\ddot{L} = \left(\sum_{i=1}^n e_i W_i - V^T V \right) / \sigma^2$$

represents intrinsic curvature of F at $\hat{\theta}$. Since this curvature is assumed to be negligible, $\ddot{L} = -V^T V / \sigma^2$ and therefore

$$\Delta_1 = \begin{pmatrix} {}^{\#}(V_1^T V_1)^{\#1} V_1^T V_2 \\ I \end{pmatrix} = \begin{pmatrix} {}^{\#}R_{11}^{\#1} & R_{12} \\ I \end{pmatrix} \quad (A.3)$$

where $V = (V_1, V_2)$ and R_{1j} is defined in (15).

An expression for Δ_2 can be obtained similarly by taking second partial derivatives of (A.1) with respect to θ_{2r} and θ_{2s} , $r, s=1, 2, \dots, p_2$. This yields

$$\sum_{b=1}^{p_1} L_{ab} \frac{\partial^2 \alpha_b}{\partial \theta_{2r} \partial \theta_{2s}} = \sum_{b=1}^p \sum_{c=1}^p L_{abc} \frac{\partial \alpha_c}{\partial \theta_{2r}} \frac{\partial \alpha_b}{\partial \theta_{2s}} \quad (A.4)$$

where L_{ab} , L_{abc} and α_b denote the indicated elements of L , L and $\alpha^T = (g^T(\theta_2), \theta_2^T)$, respectively, and $a=1, 2, \dots, p_1$. The component

$\partial^2 \alpha_b / \partial \theta_{2r} \partial \theta_{2s}$ is the (r, s) -th element of the b -th face Δ_{2b} of Δ_2 . Since $\Delta_{2b} = 0$ for $b=p_1+1, \dots, p$ the summation on the left of (A.4) need only range

from 1 to p_1 . Notice also that $\partial \alpha_c / \partial \theta_{2r}$ is simply the (c,r) -th element of Δ_1 .

Expressing (A.4) in matrix notation and solving for Δ_2 gives

$$\Delta_2 = - \begin{bmatrix} L_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta_1^T & \ddots & \Delta_1 \end{bmatrix} \quad (\text{A.5})$$

Here and in what follows brackets $[] []$ indicate column multiplication as defined in Bates and Watts (1980). (Generally, if A is an $a \times b$ matrix and B is a $b \times c \times d$ array then the elements of the i -th face C_i , $i=1, \dots, a$, of the $a \times c \times d$ array $C = [A][B]$ are $A_i^T B_{jk}$, $j=1, 2, \dots, c$, $k=1, 2, \dots, d$, where A_i^T is the i -th row of A and B_{jk} is the jk -th column of B .) As before we will take

$$L_{11} = -V_1^T V_1 / \sigma^2.$$

To further evaluate Δ_2 , we require the $p \times p \times p$ array \ddot{L} . Straightforward algebra will verify that

$$L_{abc} = \frac{1}{\sigma^2} \sum_{i=1}^n (e_i f_i^{abc} - f_i^a f_i^{bc} - f_i^b f_i^{ac} - f_i^c f_i^{ab})$$

Using this representation it is easily verified that the a -th face \ddot{L}_a of \ddot{L} is

$$\ddot{L}_a = - \frac{1}{\sigma^2} \{ [b_a^T V^T][W] + V^T K_a + K_a^T V \} \quad (\text{A.6})$$

where b_a is the a -th standard basis vector for R^p and $K_a = b_a^T W$ is the $n \times p$ matrix with W_{ac} as the c -th column. Finally, it follows from (A.6) that

$$Z^T \ddot{L} Z = - \frac{1}{\sigma^2} \{ Z^T [V^T][W]Z + 2[Z^T V^T][W]Z \} \quad (\text{A.7})$$

where Z is an arbitrary $p \times 1$ vector. This form will be useful in later developments.

A.2 Tangent plane, Term (19a)

It follows immediately from (A.3) that

$$V\Delta_1 = P_{V_1}' V_2 = U_2 R_{22} \quad (A.8)$$

where U_2 is defined following (18). Thus, the relevant tangent plane is the affine subspace $F(\hat{\theta}) + C(P_{V_1}' V_2)$. Transforming term (14a) according to (16) and (17) immediately gives term (19a).

A.3 Parameter-Effects, Term (19b)

From the form of Δ_2 given by (A.3), it is clear that term (14c) is in $C(V_1)$ and is thus orthogonal to the θ_2 -subspace tangent plane. The parameter-effects component of (14) must therefore come from term (14b).

The three-dimensional array W in (14b) can be decomposed into the sum of three arrays with orthogonal columns,

$$W = [P_V P_{V_1}'] [W] + [P_{V_1}] [W] + [P_V'] [W] \quad (A.9)$$

The first term in this decomposition contains the projections of the columns of W onto $C(P_{V_1}' V_2)$ and thus it represents parameter-effects curvature for the subset problem. The second and third terms are intrinsic components for h and F , respectively. Since the intrinsic curvature of F at $\hat{\theta}$ is assumed to be negligible, the third term of (A.9) is set to zero. Addend (14b) can now be reexpressed as

$$\frac{1}{2} \phi_2^T \Delta_1^T W \Delta_1 \phi_2 = \frac{1}{2} \phi_2^T \Delta_1^T [P_V^{-1} P_{V_1}] [W] \Delta_1 \phi_2 \quad (\text{A.10a})$$

$$+ \frac{1}{2} \phi_2^T \Delta_1^T [P_{V_1}] [W] \Delta_1 \phi_2 \quad (\text{A.10b})$$

From (18) and (A.3) it follows that

$$\tilde{W}_{22} = R_{22}^T \Delta_1^T W \Delta_1 R_{22}^{-1}$$

Using this in combination with (17) and (A.8) to transform the coordinates in term (A.10a) gives term (19b).

A.4 Intrinsic Curvature, Term (19c)

In the expansion of h given in (14), we still have the sum of terms (14c) and (A.10b) to deal with. We first consider (14c).

Using (A.5) and (A.7) with $Z = V_1 \phi_2$ we have

$$\begin{aligned} \frac{1}{2} V(\phi_2^T \Delta_2 \phi_2) &= \frac{g^2}{2} M (\phi_2^T \Delta_1^T \ddot{L} \Delta_1 \phi_2) \\ &= -\frac{1}{2} M \{ \phi_2^T \Delta_1^T [V^T] [W] \Delta_1 \phi_2 \} \\ &= M [\phi_2^T \Delta_1^T V^T] [W] \Delta_1 \phi_2 \end{aligned} \quad (\text{A.11})$$

where $M = (V_1 (V_1^T V_1)^{-1}, 0)$. The first term of (A.11) is exactly the negative of term (A.10b) so that in an obvious notation

$$\begin{aligned} (14c) + (A.10b) &= M [\phi_2^T \Delta_1^T V^T] [W] \Delta_1 \phi_2 \\ &= [\phi_2^T \Delta_1^T V^T] [M W \Delta_1] \phi_2 \end{aligned} \quad (\text{A.12})$$

From (18) and the definition of \tilde{W} , it can be shown that

$$M W \Delta_1 = U_1 \tilde{W}_{12} R_{22}$$

Finally, using this relationship, (A.8) and (17) to transform the coordinates in (A.12) we obtain term (19c).

TABLE 1

Subset Curvatures

Model/Data	Parameter Subset	r_s^{τ}	r_s^{η}	r_s
(33)	θ_1	.330	.183	.377
Bates & Watts	θ_2	.393	.089	.403
(33)	θ_1	.014	.025	.029
Michaelis & Menton	θ_2	.050	.019	.053
(34)	θ_1	.277	.148	.314
Draper & Smith	θ_2	.053	.044	.069
	θ_1	.165	.180	.244
(35)	θ_2	.003	.059	.059
Ratkowsky	θ_3	.153	.132	.203
	(θ_1, θ_3)	1.07	.088	1.07
	(θ_2, θ_3)	.518	.000	.518
	θ_1	1.75	.190	1.76
(35)	θ_2	1.80	.256	1.82
Hunt	θ_3	.018	.094	.095
	(θ_2, θ_3)	36.4	.000	36.4

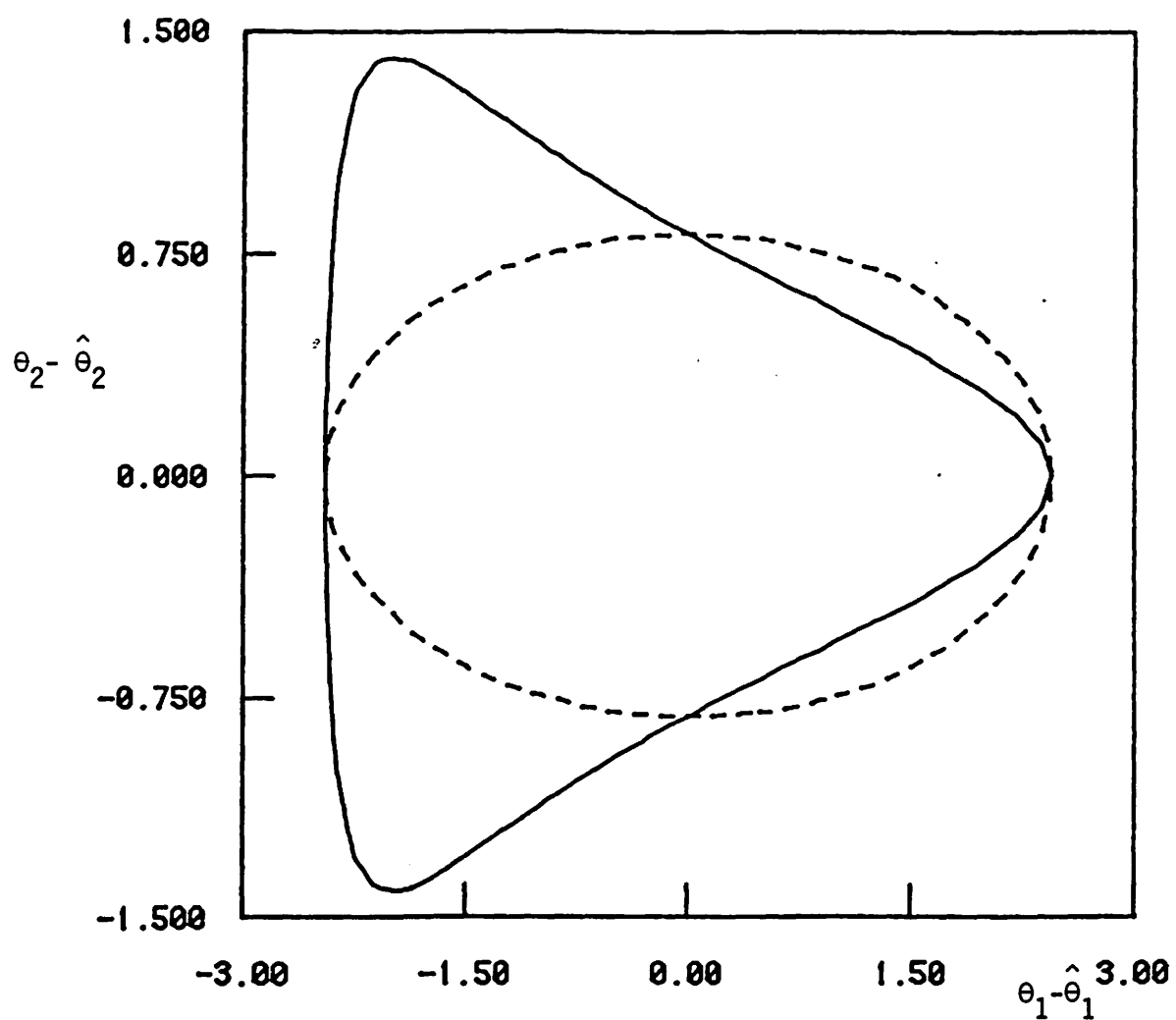


Figure 1. 95% confidence regions for (θ_1, θ_2) from the Fieller-Creasy model (8); $(\hat{\theta}_1, \hat{\theta}_2) = (3, 0)$. Likelihood region — . Tangent plane region ---- .

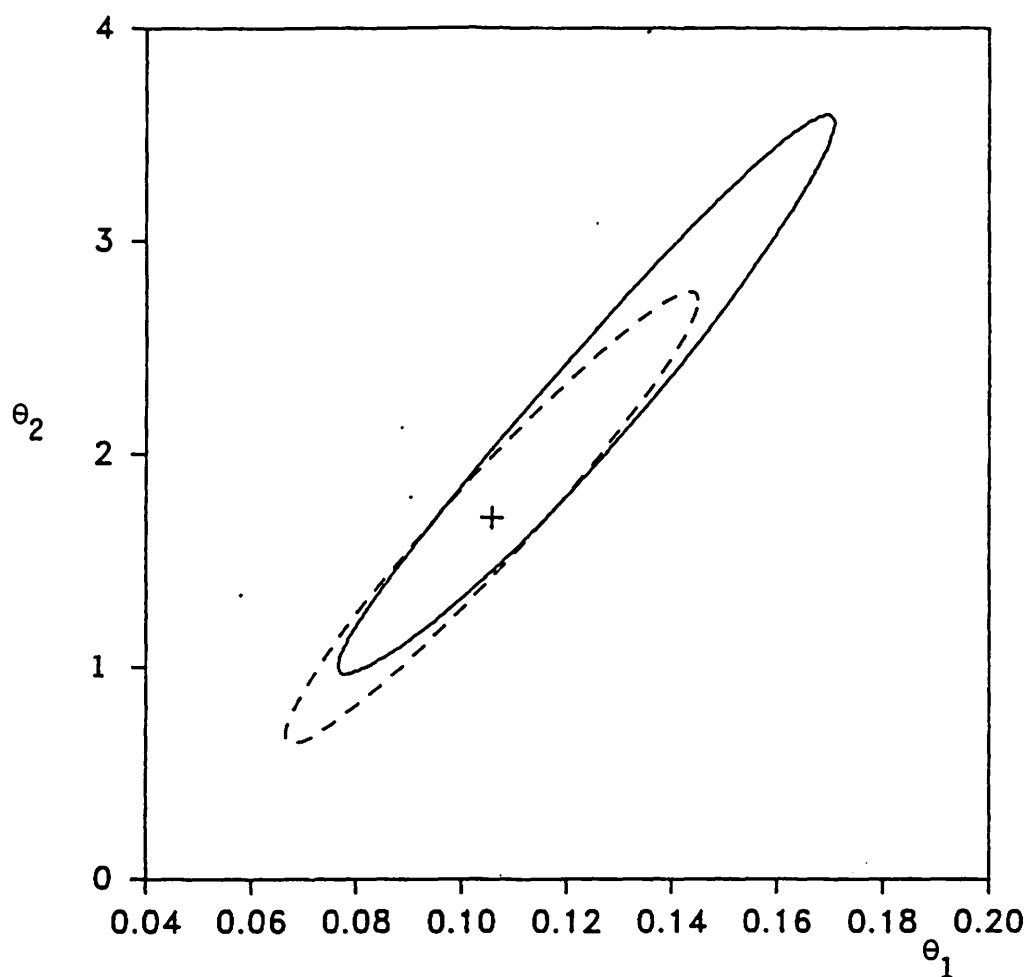


Figure 2. Nominal 87% bivariate confidence regions with 95% marginal regions for (θ_1, θ_2) from model (33) and the Bates-Watts (1980) data. Likelihood — . Tangent plane ---- .

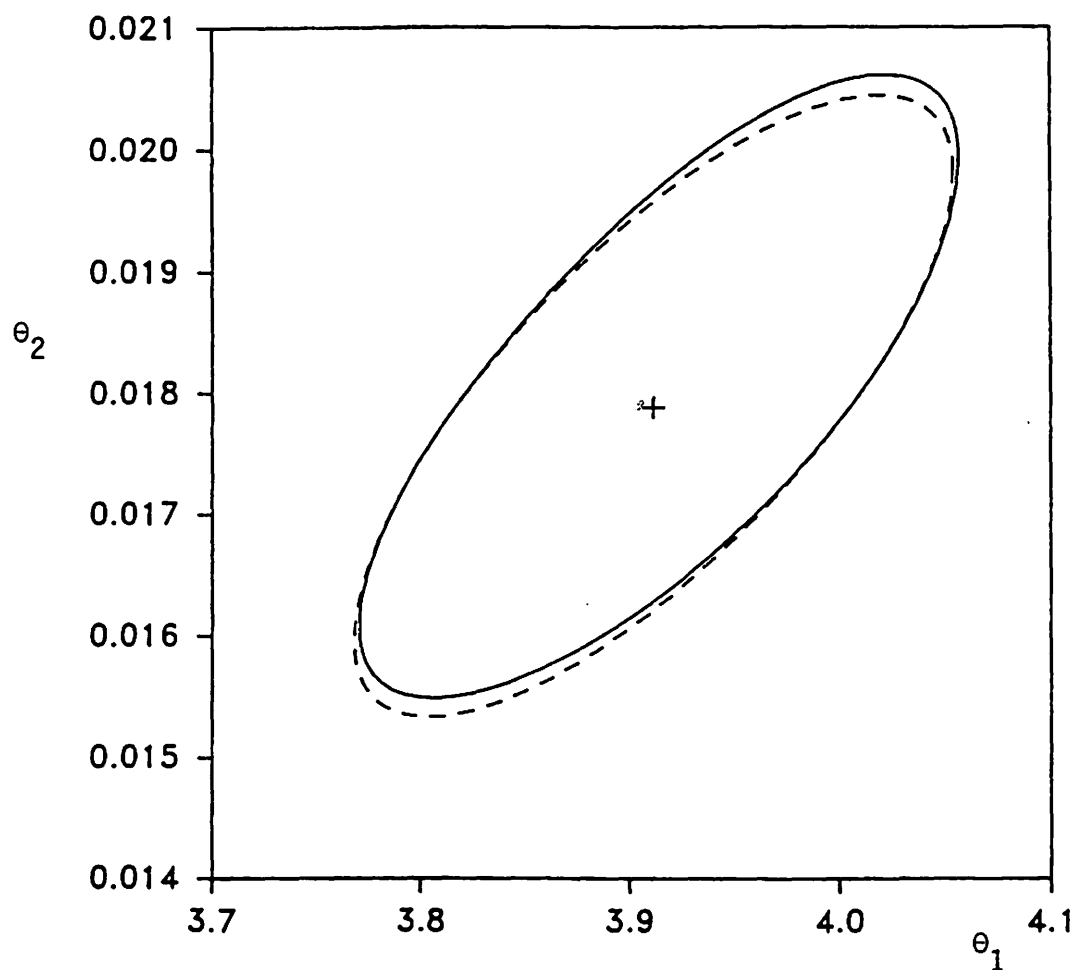


Figure 3. Nominal 88% bivariate regions with 95% marginal regions for (θ_1, θ_2) from model (33) and the Michaelis-Menton (1913) data. Likelihood — . Tangent plane ---- .

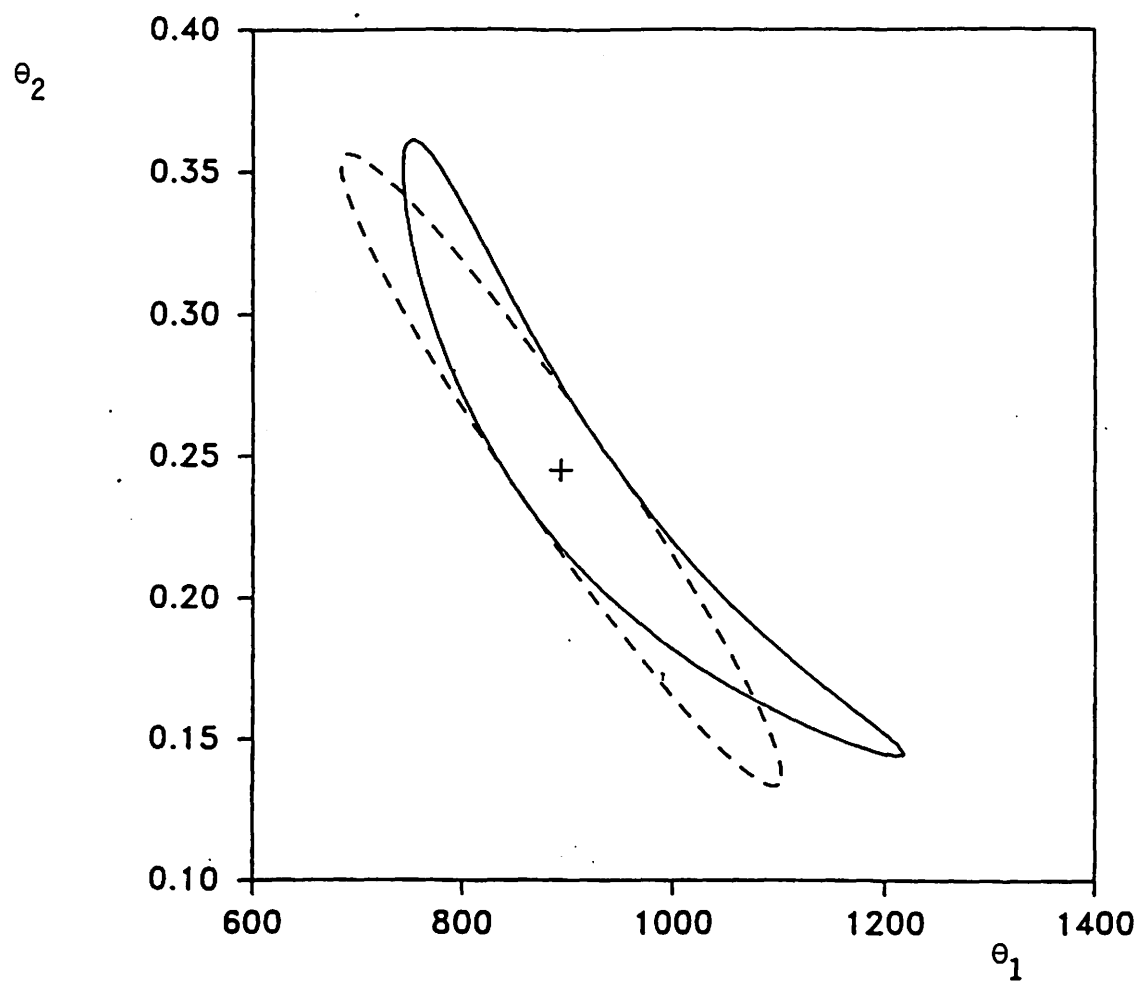


Figure 4. Nominal 88% bivariate regions with 95% marginal regions for (θ_1, θ_2) from model (34) and the Draper-Smith (1981) data. Likelihood — . Exact ---- .

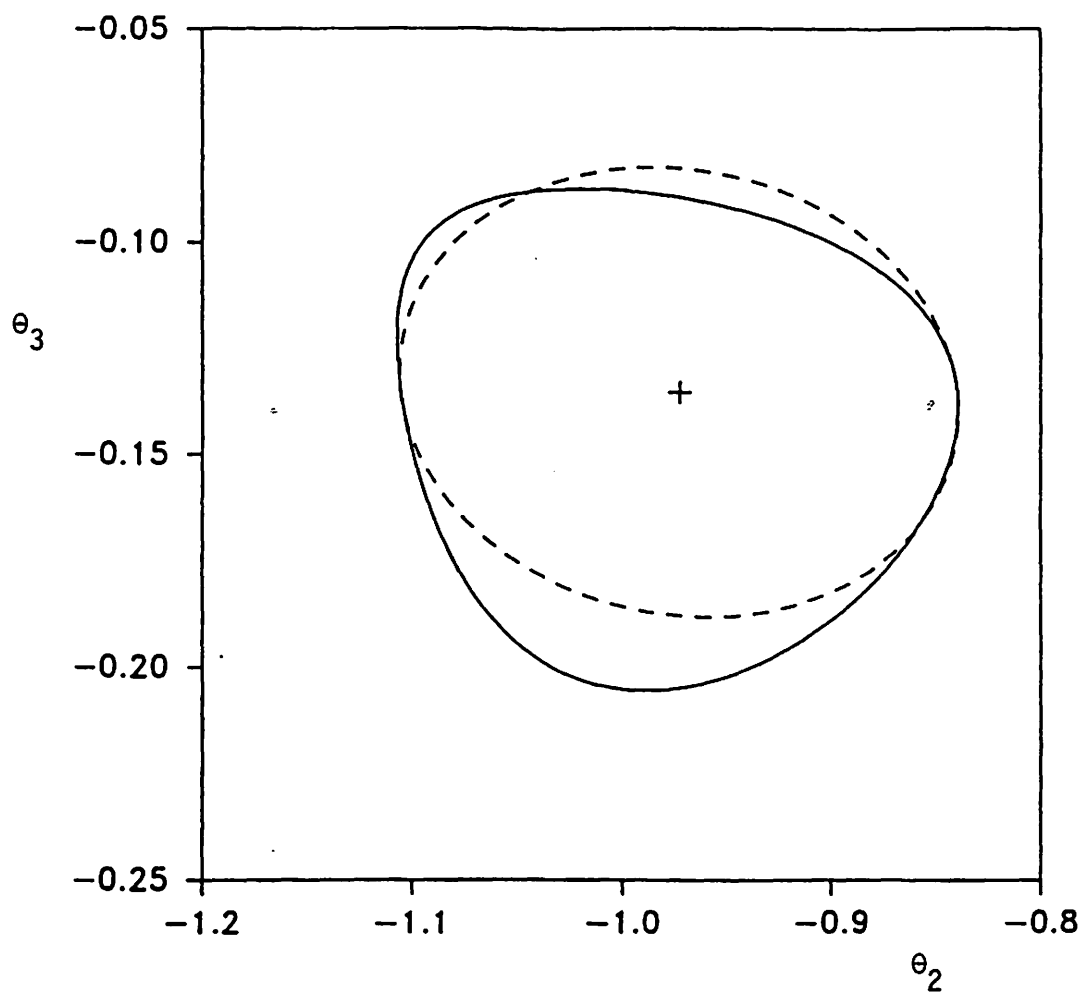


Figure 5. Nominal 86% bivariate regions with 95% marginal regions for (θ_2, θ_3) from model (35) and the Ratkowsky (1983) data. Likelihood — . Tangent plane ---- .

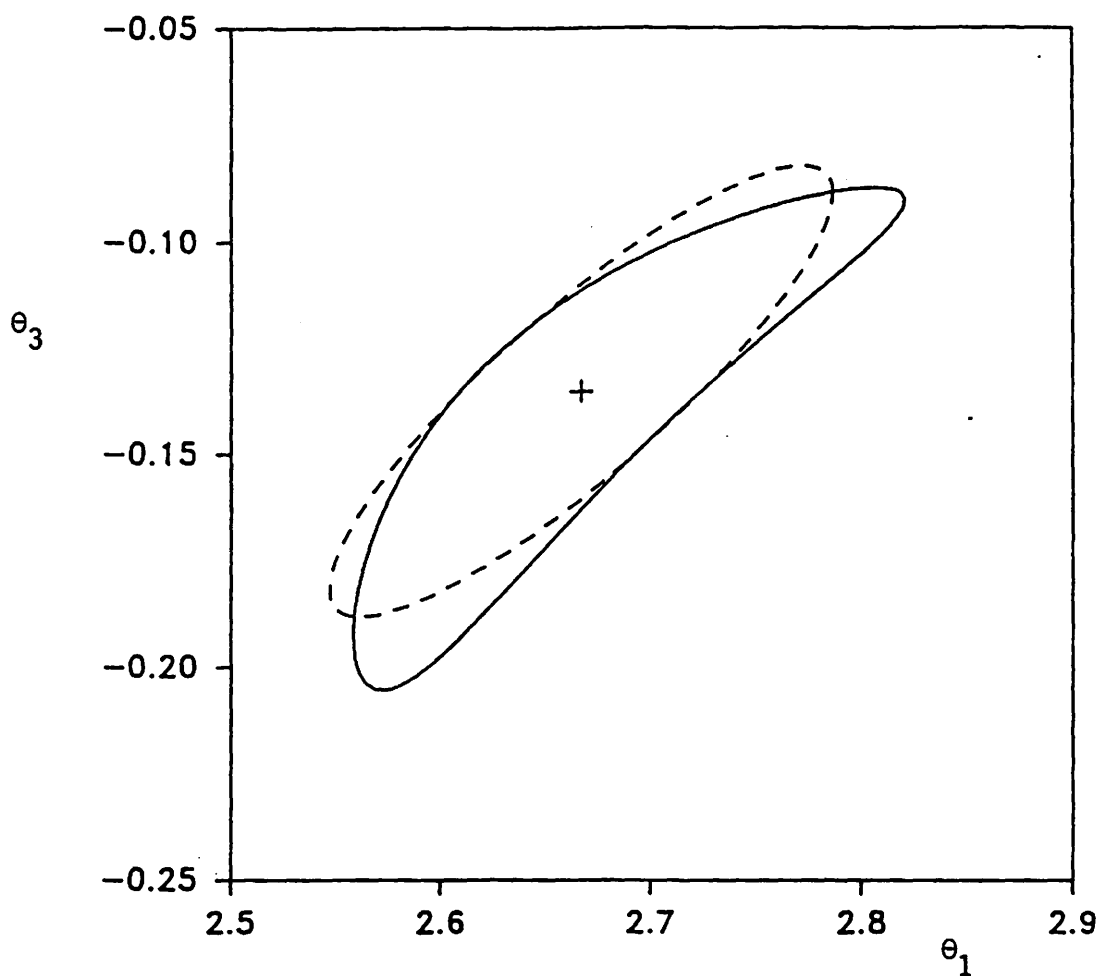


Figure 6. Nominal 86% bivariate regions with 95% marginal regions for (θ_1, θ_3) from model (35) and the Ratkowsky (1983) data. Likelihood — . Exact ---- .

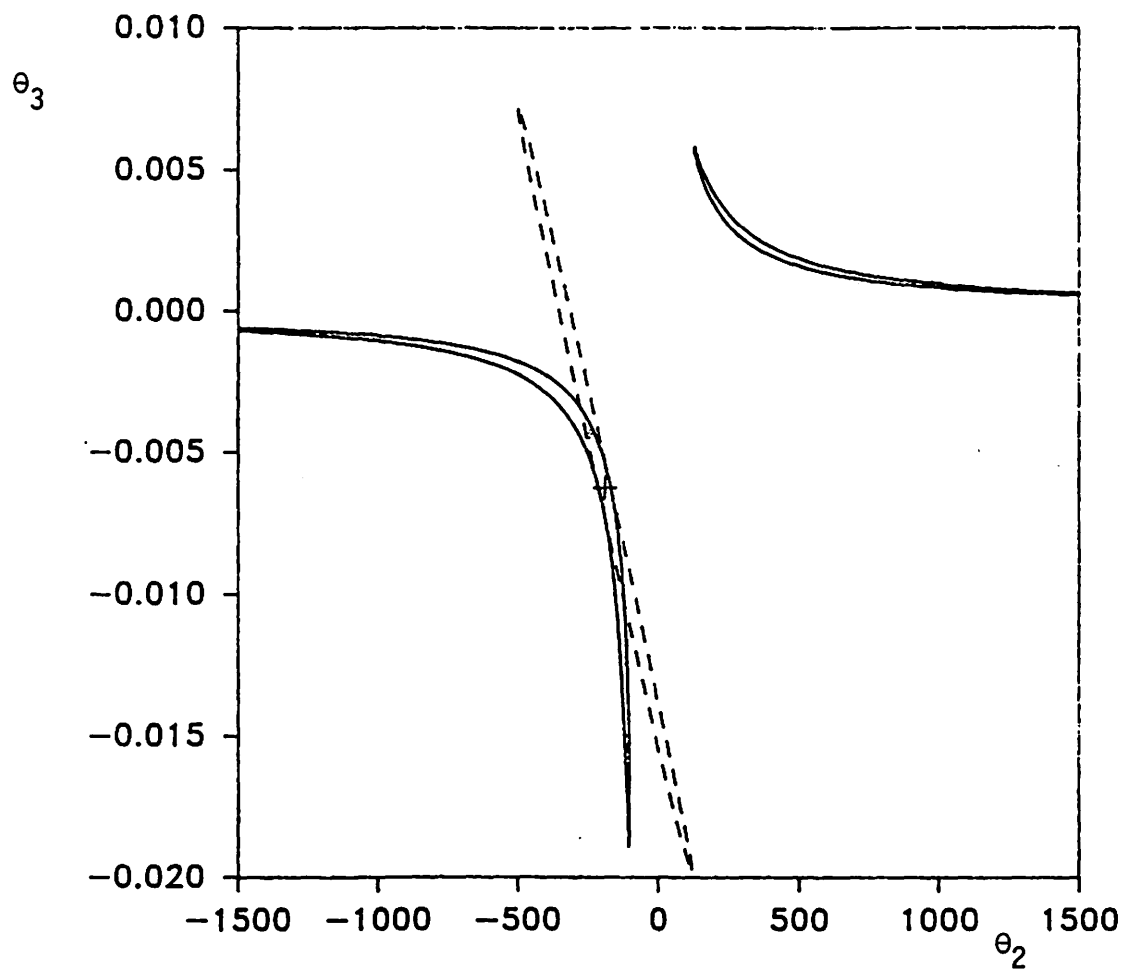


Figure 7. Nominal 87% bivariate regions with 95% marginal regions for (θ_2, θ_3) from model (35) and the Hunt (1970) data. Likelihood — . Exact ---- .

REFERENCES

- Bates, D. M. and Watts, D. G. (1980), "Relative Curvature Measures of Nonlinearity," Journal of the Royal Statistical Society, Ser. B, 42, 1-25.
- Bates, D. M. and Watts, D. G. (1981), "Parameter Transformations for Improved Approximate Confidence Regions in Nonlinear Least Squares", The Annals of Statistics, 9, 1152-1167.
- Bates, D. M., Hamilton, D. C. and Watts, D. G. (1983), "Calculation of Intrinsic and Parameter-Effects Curvatures for Nonlinear Regression Models", Communications in Statistics, Part B-Simulation and Computation, 12, 469-477.
- Cook, R. D. and Witmer, J. A. (1984), "A Note on Parameter-effects Curvature", Technical Report 439, School of Statistics, University of Minnesota.
- Cox, R. D. and Hinkley, D. V., (1974), Theoretical Statistics, London: Chapman and Hall.
- Dongarra, J. J., Moler, C. B., Bunch, J. R., and Stewart, G. W. (1979), LINPACK Users' Guide, Philadelphia: SIAM.
- Draper, N. R. and Smith, H. (1981), Applied Regression Analysis, New York: John Wiley.
- Goldberg, M. L., Bates, D. M. and Watts, D. G. (1984), "Curvature Measures Avoiding Second Derivatives", Technical Report No. 739, Department of Statistics, University of Wisconsin.
- Hamilton, D. C., Watts, D. G. and Bates, D. M. (1982), "Accounting for Intrinsic Nonlinearity in Nonlinear Regression Parameter Inference Regions", The Annals of Statistics, 10, 386-393.

Hunt, W. F. (1970), "The Influence of Leaf Death on the Rate of Accumulation of Green Herbage During Pasture", Journal of Applied Ecology, 7, 41-50.

Kennedy, W. J. and Gentle, J. E. (1980), "Statistical Computing", New York: Marcel Dekker.

Linssen, H. N. (1980), "Discussion of the Paper by Dr. Bates and Professor Watts", Journal of the Royal Statistical Society B, 42, p. 21.

Michaelis, L. and Menton, M. L. (1913), "Kinetik der Invertinwirkung", Biochemische Zeitschrift, 49, 333.

Ratkowsky, D. A. (1983). Nonlinear Regression Modeling. Marcel Dekker, Inc.: New York.